Genet Resour Crop Evol (2009) 56:311–321
DOI 10.1007/s10722-008-9365-3

RESEARCH ARTICLE

Cloning and phylogenetic analysis of polyphenol oxidase genes in common wheat and related species

X. Y. He · Z. H. He · C. F. Morris · X. C. Xia

Received: 7 March 2008 / Accepted: 15 July 2008 / Published online: 9 August 2008

Abstract Cloning and phylogenetic analysis of polyphenol oxidase (PPO) genes in common wheat and its relatives would greatly advance the understanding of molecular mechanisms of grain PPO activity. In the present study, six wheat relative species, including *T. urartu*, *T. boeoticum*, *T. monococcum*, *T. dicoccoides*, *T. durum* and *Ae. tauschii*, were sampled to isolate new alleles at *Ppo-A1* and *Ppo-D1* loci corresponding to common wheat PPO genes, and seven new alleles were identified from these species, which were designated as *Ppo-A1c* (from *T. urartu*), *Ppo-A1d* (*T. boeoticum*), *Ppo-A1e* (*T. monococcum* and *T. durum*), *Ppo-A1f* (*T. dicoccoides*), *Ppo-A1g* (*T. durum*), *Ppo-D1c* (*Ae. tauschii*) and *Ppo-D1d* (*Ae. tauschii*), respectively. Five out of the seven alleles detected in the wheat relatives contained an open reading frame (ORF) of 1,731 bp, encoding a polypeptide of 577 residues, which is the same as those of *Ppo-A1* and *Ppo-D1* genes in common wheat, whereas, the full-length ORF of the allele *Ppo-A1g* from *T. durum* was not obtained, and a 73-bp deletion occurred in the third exon of *Ppo-D1d*, an allele from *Ae. tauschii*, resulting in a shorter polypeptide of 466 amino acids. The 191-bp insertion in the first intron reported previously in common wheat was also found in *T. dicoccoides* lines, implying that more than one tetraploid wheat lines may be involved in the origination of common wheat. Phylogenetic trees were constructed using the genomic DNA sequences of the seven alleles, together with four from common wheat and four partial PPO gene sequences deposited in GenBank. The genome tribe A was divided into two clusters, one of which contained *Ppo-A1d* and *Ppo-A1e*, and the other included the remaining five alleles at *Ppo-A1* locus. The alleles from different clusters showed high sequence divergences, indicated by dozens of SNPs and five to six InDels. The genome tribe D comprised the alleles *Ppo-D1a*, *Ppo-D1c*, *Ppo-D1d* and *Ppo-D1b*, and the former three were clustered together, showing significant sequence divergence from *Ppo-D1b*. In addition, the relationships between these allelic variants and grain PPO activities were also discussed.

Electronic supplementary material The online version of this article (doi:10.1007/s10722-008-9365-3) contains supplementary material, which is available to authorized users.

X. Y. He · Z. H. He (✉) · X. C. Xia (✉)
Institute of Crop Science, National Wheat Improvement Center/The National Key Facility for Crop Gene Resources and Genetic Improvement, Chinese Academy of Agricultural Sciences (CAAS), Zhongguancun South Street 12, Beijing 100081, China
e-mail: zhhe@public3.bta.net.cn

X. C. Xia
e-mail: xiaxianchun@caas.net.cn

Z. H. He
International Maize and Wheat Improvement Center (CIMMYT) China Office, c/o CAAS, Zhongguancun South Street 12, Beijing 100081, China

C. F. Morris
USDA-ARS Western Wheat Quality Laboratory, Washington State University, P.O. Box 646394, Pullman, WA 99164-6394, USA

Keywords Common wheat · Grain PPO activity · Phylogenetics · Polyphenol oxidase (PPO) genes · *Triticum aestivum* · Wheat relatives

Introduction

Common wheat (*Triticum aestivum* L.) is an allohexaploid species, constituted of A, B and D genomes. Its origination and evolution have been investigated extensively, and it is generally accepted that two evolutionary events contributed to the formation of hexaploid common wheat (Feuillet et al. 2001; Huang et al. 2002; Gu et al. 2004; Petersen et al. 2006). The first one was the hybridization between *T. urartu* Thum. (the A genome donor) and *Ae. speltoides* Tausch or a closely related species (the B genome donor), resulting in the formation of *T. dicoccoides* (Körn. ex Asch. et Graeb.) Schweinf. (wild emmer wheat, AABB). Subsequently, *T. dicoccon* Schrank (emmer, a domesticated form of *T. dicoccoides*) hybridized with *Ae. tauschii* Cosson (the D genome donor), producing hexaploid common wheat (AABBDD). Durum wheat (*T. durum* Desf., AABB), an important cereal used for making pasta, is another domesticated form of *T. dicoccoides*, and closely related to *T. dicoccon* (Salamini et al. 2002; Ozkan et al. 2005; Luo et al. 2007; Jauhar 2007). *T. boeoticum* Boiss. (wild einkorn wheat, A^mA^m) is a wild diploid wheat species closely related to *T. urartu*, and its domesticated form *T. monococcum* L. is still being cultivated to a limited extent (Gill and Friebe 2002; Salamini et al. 2002).

Polyphenol oxidases (PPOs) are copper-containing metalloenzymes, which are encoded in the nucleus and then transported into the plastids (Anderson and Morris 2003). PPOs are able to catalyze the oxidation of phenols to produce quinones that subsequently react with amino acids and proteins to form brown and black pigments, greatly reducing the appearance quality of wheat products (Feillet et al. 2000). It has been proven that PPOs are main causal factors involved in darkening discoloration of wheat-based products, such as steamed bread (Dexter et al. 1984), pan bread (McCallum and Walker 1990), pasta (Simeone et al. 2002), and especially Asian noodles (Kruger et al. 1992; Baik et al. 1995; Mares and Campbell 2001; Fuerst et al. 2006). Although no evidences have been proposed that high PPO activity is associated with decreased nutritional value, the darkened color can still negatively affect consumer's choice (Simeone et al. 2002). Thus, development of wheat cultivars with low grain PPO activity is one of the main objectives in wheat breeding programs (Mares and Campbell 2001; Jukanti et al. 2004).

In many plant species, such as tomato (Newmann et al. 1993), potato (Thygesen et al. 1995), apple (Boss et al. 1995), banana (Gooding et al. 2001) and red clover (Sullivan et al. 2004), PPO genes are organized in multigenic families. In common wheat, Kruger (1976) reported 12 isozymes of PPO, implying that PPOs are encoded by a multi-gene family. Jukanti et al. (2004) cloned six partial wheat PPO genes and classified them into two clusters with each three genes, and one cluster was proposed to be involved in the developing kernels and the other in the leaves and stems. Recently, Massa et al. (2007) cloned 21 partial PPO gene sequences from kernels in wheat related species, and phylogenetic analyses indicated that the genes could be classified into four major clusters. Based on the fact that three PPO genes falling into different clusters were isolated from *T. monococcum*, Massa et al. (2007) proposed that gene duplication events happened in PPO genes and led to the formation of multigenic families. Furthermore, Fuerst et al. (2008) demonstrated the presence of two or more kernel-type PPO genes in diploid progenitors of common wheat A, B and D genomes by southern analysis, providing more evidences for the multigenic organization of PPO genes in wheat related species.

The role of PPOs has not been conclusively delineated. However, the enzymes are often ascribed to plant defense (Constabel et al. 2000; Haruta et al. 2001). In this regard, it is relevant to note that two distinct PPO gene families have been identified, each with apparently exclusive tissue specificity (Jukanti et al. 2004, 2006). Consequently, it may be a reasonable breeding objective to eliminate PPO activity from the grain so that it does not exert negative effects on food end-product color.

Grain PPO activity in common wheat has been studied extensively (Kruger et al. 1994; Jimenez and Dubcovsky 1999; Anderson and Morris 2001; Demeke et al. 2001; Fuerst et al. 2006). Many reports indicate that the PPO genes on homoeologous group 2 chromosomes are responsible for grain PPO activity, especially those on chromosomes 2A and 2D (Jimenez and Dubcovsky 1999; Anderson and Morris 2001;

Demeke et al. 2001; Mares and Campbell 2001; Simeone et al. 2002; Sun et al. 2005; Raman et al. 2005; He et al. 2007). In a few studies, a QTL with minor effect was detected on chromosome 2B (Demeke et al. 2001; Watanabe et al. 2004). Previously we cloned full coding sequences of the PPO genes on chromosomes 2A (*Ppo-A1*) and 2D (*Ppo-D1*) in common wheat, and identified two allelic variants at each of the two loci. All of these genes have an ORF of 1,731 bp, and comprise three exons and two introns (He et al. 2007). To the best of our knowledge, apart from several partial sequences (Massa et al. 2007), no full-length coding sequences of PPO genes have been isolated from relatives of common wheat so far.

Considering the significance of *Ppo-A1* and *Ppo-D1* for grain PPO activity, cloning orthologs of the two genes in the species related to common wheat would essentially advance our knowledge on the molecular mechanisms of grain PPO activity. The objectives of this study were to isolate new alleles of *Ppo-A1* and *Ppo-D1* from the species related to common wheat, and to analyze their sequence characteristics and phylogenetic relationships with the PPO genes in common wheat.

Materials and methods

Plant materials

Five accessions of *T. boeoticum*, four of *T. monococcum*, one of *T. urartu*, five of *T. dicoccoides* and 16 of *T. durum* were used for cloning the *Ppo-A1* gene, while 15 accessions of *Ae. tauschii* were employed to clone the *Ppo-D1* gene (Supplementary Table 1). All these accessions were provided by National Key Facilities for Crop Genetic Resources and Improvement (NFCRI), Institute of Crop Science, CAAS, China.

Strategies for identification of new PPO alleles in common wheat relatives

Genomic DNA was isolated from single kernels following the method modified from Lagudah et al. (1991). To prevent the admixture of germplasm, three morphologically representative seeds from each accession were subjected to DNA extraction, and once the three seeds from an accession exhibited inconsistent genotypes, more seeds would be analyzed to determine the representative one for the accession.

New PPO alleles were detected in the plant materials with the PPO markers published previously, with extensive polymorphisms among common wheat cultivars. For diploid and tetraploid wheat species, *PPO18* (Sun et al. 2005) were used to detect new *Ppo-A1* alleles, and its PCR products were directly sequenced using the upstream primer of the marker, with three seeds analyzed for each accession. Sequenced PCR products were compared with the two alleles identified in common wheat (He et al. 2007), *Ppo-A1a* and *Ppo-A1b*, to find if the sequences represent new *Ppo-A1* alleles. As for *Ae. tauschii* lines, markers *PPO16*/*PPO29* (He et al. 2007) and *WP3-2* (Chang et al. 2007) were used to detect new *Ppo-D1* alleles, and the PCR products were directly sequenced with the upstream primers of the markers and compared with *Ppo-D1a* and *Ppo-D1b*, two alleles at common wheat *Ppo-D1* locus (He et al. 2007). Gene sequencing were performed in Shanghai Sangon Biological Engineering & Technology and Service Co., Ltd. (http://www.sangon.com).

Based on the results of genotyping, the accessions were classified into different groups, and representative accessions with one from each group were chosen for cloning full-length PPO alleles. Two primer sets, *PPO1* and *PPO2*, were designed according to the common wheat PPO gene sequence of *Ppo-A1a* (EF070147) to isolate its orthologs in diploid and tetraploid wheat lines, and two primer sets, *PPO3* and *PPO4*, were developed based on *Ppo-D1a* (EF070149) to clone its orthologs in *Ae. tauschii* lines. To prevent the formation of chimeric products, the amplified regions of *PPO1* and *PPO2* have an overlap of 270 bp (counted from *Ppo-A1a*), including the first intron of the PPO gene, while those of *PPO3* and *PPO4* have an overlap of 436 bp (counted from *Ppo-D1a*), covering the second intron. Primers were designed using the software Premier Primer 5 (http://www.premierbiosoft.com) and synthesized by Beijing Augct Biological Technology Co., Ltd. (http://www.augct.com).

PCR reactions were performed in an MJ Research PTC-200 thermal cycler in a total volume of 20 μl including 20 mM of Tris–HCl (pH 8.4), 20 mM of KCl, 150 μM of each of dNTPs, 1.5 mM of $MgCl_2$,

8 pmol of each primer, 1 unit of *Taq* DNA polymerase (TIANGEN Biotech Co., Ltd., www.tiangen.com) and 50 ng of genomic DNA. Reaction conditions were 95°C for 5 min, followed by 40 cycles of 95°C for 30 s, 62–65°C for 30 s (according to the annealing temperatures of different primer sets), and 72°C for 1.5 min, with a final extension of 72°C for 5 min. The PCR products were separated in a 1.5% agarose gel, stained with ethidium bromide, and visualized using UV light. The targeted bands were recovered and cloned into pMD18-T vector (TaKaRa Biotechnology (Dalian) Co., Ltd.) following the manufacturer's instructions. To eliminate errors in sequencing, the PCR reaction and DNA sequencing were repeated 2–4 times for each primer set.

Complete genomic DNA sequences of PPO genes were constructed either by the PCR products of *PPO1* and *PPO2* (in diploid and tetraploid wheat lines), or those of *PPO3* and *PPO4* (in *Ae. tauschii* lines). Intron positions were determined by the alignment of amplified genomic DNA sequences and their corresponding orthologs in common wheat, using the software DNAMAN (http://www.lynnon.com).

Phylogenetic analyses

All the genomic DNA sequences of cloned PPO genes in this study, together with those of common wheat genes *Ppo-A1a*, *Ppo-A1b*, *Ppo-D1a* and *Ppo-D1b* (He et al. 2007), and four partial PPO gene sequences registered in GenBank (accession numbers DQ889693, DQ889700, DQ889706 and DQ889708, Massa et al. 2007) were used to construct the phylogenetic trees. Gene sequences were aligned with the software ClustalW 1.83 (Thompson et al. 1997). Neighbor joining tree and Maximum parsimony tree were generated by the program MEGA version 3.1 (Kumar et al. 2004) with 1,000 replicates of bootstrap testing, in which gaps were excluded only from pairwise comparisons for the former algorithm, whereas all sites were used as characters for the latter one. In addition, a Maximum likelihood tree was generated by the software package PHYLIP with default parameters.

Results

Cloning of PPO genes in wheat relative species

New allelic variants at *Ppo-A1* and *Ppo-D1* loci in wheat relative species were identified with the markers *PPO18*, *PPO16*, *PPO29* and *WP3-2*. In the test with *PPO18*, two polymorphic sequences were detected in *T. dicoccoides* and two in *T. durum* (Table 1), whereas no polymorphisms were found in diploid wheat species. Markers *PPO29* and *WP3-2* showed no polymorphisms, i.e., the former yielded no PCR products and the latter amplified a 563-bp fragment in each of the *Ae. tauschii* lines. However, *PPO16* amplified a 713-bp fragment from *Ae. tauschii* Ae38, while it generated no PCR products

Table 1 Plant materials used in this study and their PPO alleles

Species	Genome	Accession[a]	PCR profile (bp)[b]	Allele	GenBank accession number
T. urartu	A^u	UR1, (1)	685	*Ppo-A1c*	EU371651
T. boeoticum	A^m	BO1, (5)	716	*Ppo-A1d*	EU371652
T. monococcum	A^m	MO1, (4)	716	*Ppo-A1e*	EU371653
T. dicoccoides	AB	DS3, (4)	685	*Ppo-A1f*	EU371654
		DS4, (1)	876	*Ppo-A1b*	EF070148
T. durum	AB	DR8, (7)	716	*Ppo-A1e*	EU371653
		Langdon, (9)	Null	*Ppo-A1g*	EU371655
Ae. tauschii	D	Ae38, (1)	713	*Ppo-D1c*	EU371656
		Y59, (14)	Null	*Ppo-D1d*	EU371657

[a] The number of accessions of each genotype was indicated in parentheses, and a representative accession was shown for each allele. Accession identifiers were from NFCRI, Institute of Crop Science, CAAS, China

[b] Diploid and tetraploid wheat lines were detected with *PPO18*, while *Ae. tauschii* lines were amplified with *PPO16*

in the remaining *Ae. tauschii* lines. Therefore, *PPO16* identified two allelic variants in the *Ae. tauschii* lines (Table 1).

Based on the polymorphic tests, seven new allelic variants were identified in these species, with five at *Ppo-A1* and two at *Ppo-D1* locus (Table 1). Using the primer sets *PPO1* and *PPO2* for *Ppo-A1*, and *PPO3* and *PPO4* for *Ppo-D1* (Table 2), we cloned full-length genomic DNA sequences of six alleles, i.e. *Ppo-A1c* (from *T. urartu*, EU371651), *Ppo-A1d* (*T. boeoticum*, EU371652), *Ppo-A1e* (*T. monococcum* and *T. durum*, EU371653), *Ppo-A1f* (*T. dicoccoides*, EU371654), *Ppo-D1c* (*Ae. tauschii*, EU371656) and *Ppo-D1d* (*Ae. tauschii*, EU371657). As for *Ppo-A1g* (EU371655), an allele from *T. durum* cv. 'Langdon', only the upstream sequence was amplified with *PPO1*, and the downstream sequence remains unknown. Although the coding sequences of these alleles showed high size divergences, ranging from 2,152 bp (*Ppo-A1b*) to 1,627 bp (*Ppo-D1d*), but they demonstrated high sequence identities ranging from 86.7% to 99.9%, reflecting their close phylogenetic relationships. Except *Ppo-A1g* and *Ppo-D1d*, all other five alleles shared the same size of ORF with 1,731 bp, encoding a polypeptide of 577 residues. In *Ppo-D1d*, however, a 73-bp sequence deletion occurred in its third exon, resulting in a frame-shift mutation, which subsequently led to a premature translation termination, producing a shorter polypeptide of 466 amino acids (Fig. 1).

Structures and polymorphisms of the introns

Apart from *Ppo-A1g*, all PPO gene alleles cloned in this study have similar exon/intron structures, i.e. three exons spaced by two introns. The first introns of the alleles comply with the GT-AG rule (Supplementary Figs. 1, 2), starting with a GT dinucleotide and ending with an AG dinucleotide; all of the second introns, however, fall into the class of non-canonical introns characterized by a GC-AG structure (Fig. 2). Compared with the second introns, the first ones exhibited higher sequence similarities except one InDel (Supplementary Fig. 1). A 191-bp insertion, reported previously in common wheat (Sun et al. 2005, He et al. 2007), was also found in *T. dicoccoides* DS4. The second introns differed from each other by a great number of SNPs and several InDels (Fig. 2), and among the seven PPO alleles cloned in this study and the four alleles cloned in common wheat (*Ppo-A1a*, *Ppo-A1b*, *Ppo-D1a* and *Ppo-D1b*, He et al. 2007), five types of divergences were found in the second introns, designated *a* through *e*, respectively (Fig. 2). A large InDel of approximately 200 bp was shown in the second introns compared with DQ889708, a partial PPO gene reported previously (Massa et al. 2007). Interestingly, footprints of a Miniature Inverted-repeat Transposable Element (MITE) from the *Stowaway* family were found in the second introns of all variants (Fig. 2).

Phylogenetic inferences and sequence comparisons

Using programs MEGA and PHYLIP, three trees were generated by different algorithms, and only the neighbor joining tree is presented here (Fig. 3). Topologies of the three trees are highly similar to each other, and they differed only in the relative position of *Ppo-D1b* and DQ889708, i.e. in the neighbor joining tree, *Ppo-D1b* is located on the inner clade closed with other alleles of *Ppo-D1*, while it

Table 2 Primer sequences used for the cloning of PPO alleles

	Primer sequence (5′-3′)[a]	Annealing temperature (°C)
PPO1	F: GCGGCAGCCAGAAAGCAA	63
	R: GACGACCTGTTGGCGTAGAT	
PPO2	F: CCAGATACACAACTGCTGGC	62
	R: GGAGACGCACCATAGATTCCA	
PPO3	F: GCAGCATGGAGAGCAGTCGC	65
	R: GTACGGCTGTCCCAGGAACA	
PPO4	F: AGGTTCTACGTCTACTTCCAC	62
	R: TAGATTCTATACCACGGCGAGC	

[a] The primer sets *PPO1* and *PPO2* were used to amplify the alleles of *Ppo-A1* from diploid and tetraploid wheat species, while *PPO3* and *PPO4* for those of *Ppo-D1* from *Ae. tauschii* lines

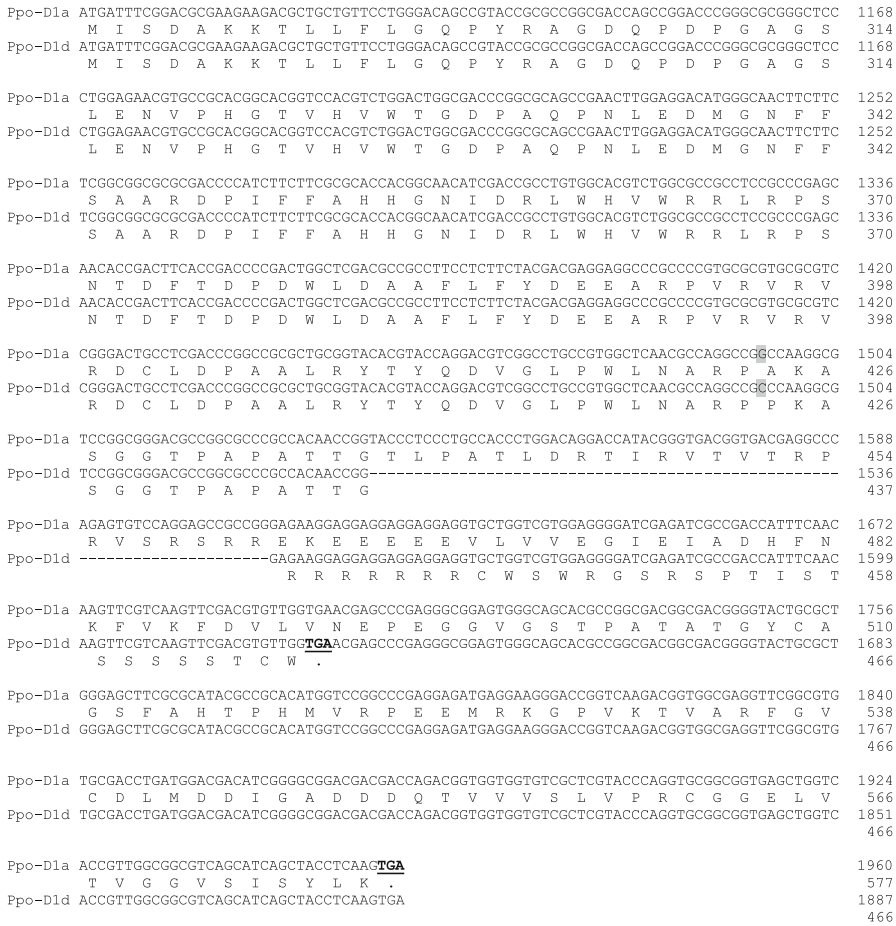

Fig. 1 Sequence alignment of the third exon of *Ppo-D1a* and *Ppo-D1d* and their deduced amino acids. Gaps are indicated with dashes, SNPs are shadowed, and the terminal codon is underlined and in bold. Sequence numbers for DNA and protein sequences were counted from the start codon and the first amino acid, respectively

Fig. 2 Alignment of the second intron from different PPO alleles. The dinucleotides at 5′ and 3′ termini are boxed, and *a* denotes the second intron of *Ppo-A1a*, *Ppo-A1b*, *Ppo-A1c* and *Ppo-A1f*, and *b* of *Ppo-A1d*, *c* of *Ppo-A1e*, *d* of *Ppo-D1a*, *Ppo-D1c* and *Ppo-D1d*, *e* of *Ppo-D1b*, *f* of DQ889708, and *g* of DQ889706. The recognition sites of a *Stowaway* element and its putative footprints are underlined and in bold

was located on the outer clade in the maximum likelihood tree and formed a cluster with DQ889708 in the maximum parsimony tree. Other clades and clusters were identical among the three trees and were well supported by high bootstrap values (Fig. 3). The genome tribe A was further divided into two major clusters (Fig. 3). Cluster I consisted of the alleles from *T. urartu* (*Ppo-A1c*), *T. dicoccoides* (*Ppo-A1b* and *Ppo-A1f*), *T. aestivum* (*Ppo-A1a* and *Ppo-A1b*), and one of the two alleles cloned from *T. durum* (*Ppo-A1g*). It is notable that *T. dicoccoides* DS4 and *T. aestivum* cv. Chinese Spring shared the same allele, *Ppo-A1b*, which was highly homologous to that of *T. durum*, *Ppo-A1g*. The allele from *T. dicoccoides* DS3 (*Ppo-A1f*) and that from *T. aestivum* cv. Nongda 139 (*Ppo-A1a*) showed high sequence identity, with only one SNP in the third exon, resulting in a shift from threonine to methionine in deduced amino acid sequences. Although the allele from *T. urartu* (*Ppo-A1c*) resided on the peripheral clade of cluster I, only several SNPs were detected between *Ppo-A1c* and other *Ppo-A1* alleles in this cluster, demonstrating their close relationships. Cluster II comprised the alleles from *T. boeoticum* (*Ppo-A1d*), *T. monococcum* (*Ppo-A1e*) and *T. durum* (*Ppo-A1e*). A same allele, *Ppo-A1e*, was detected in *T. monococcum* MO1 and *T. durum* DR8, which highly related to *Ppo-A1d*, with only two SNPs. However, the alleles in cluster II exhibited significant sequence divergences from those in cluster I, indicated by dozens of SNPs and five to six InDels (Supplementary Fig. 1). The genome tribe D included two alleles from common wheat (*Ppo-D1a* and *Ppo-D1b*), and two from *Ae. tauschii* (*Ppo-D1c* and *Ppo-D1d*). *Ppo-D1a* is highly similar to *Ppo-D1c*, an allele from *Ae. tauschii* Ae38, with only one SNP in the first exon, resulting in a mutation from serine to asparagine in deduced amino acid sequences. *Ppo-D1b* is markedly different from other alleles at *Ppo-D1* locus, as indicated by its location on the most outer clade of the tribe. Three partial PPO gene sequences, DQ889693, DQ889700 and DQ889706, formed an

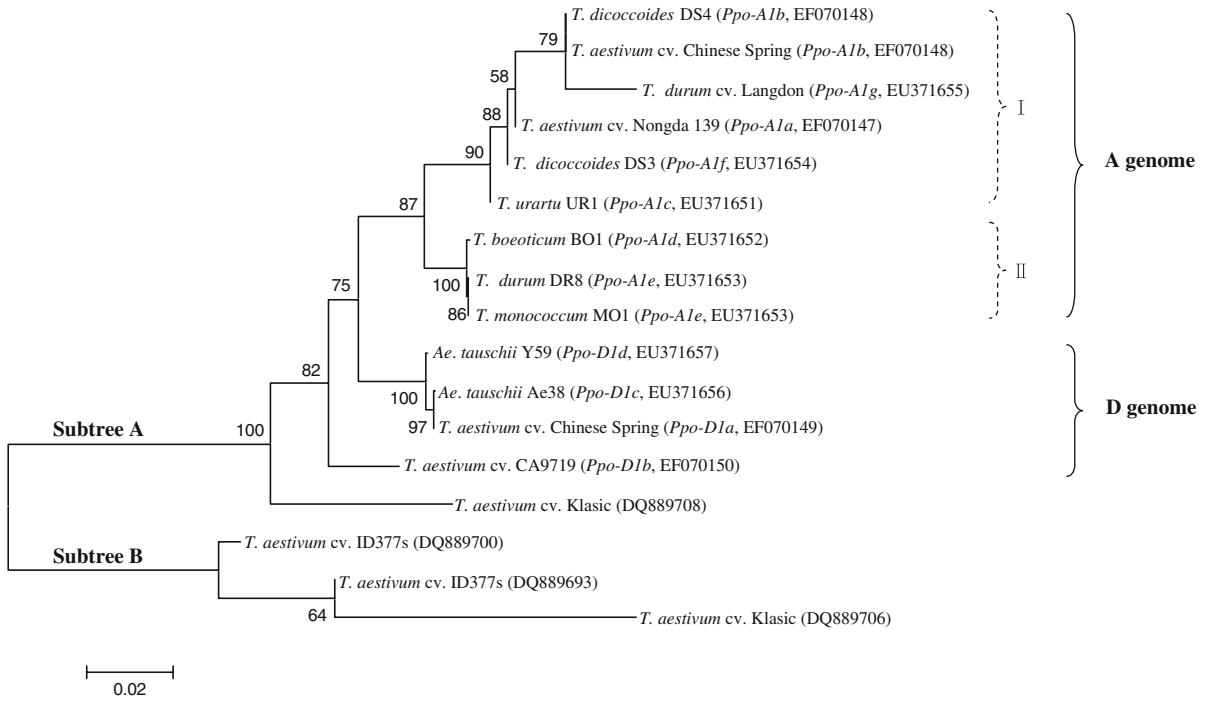

Fig. 3 Phylogenetic tree of the PPO gene families in common wheat and related species. The tree was constructed by the software MEGA version 3.1 with neighbor joining algorithm, including seven new alleles found in this study and four alleles in common wheat reported previously (He et al. 2007), and four partial PPO gene sequences deposited in GenBank (Massa et al. 2007). Genes are labeled by species name, accession identifier, allele name and GenBank accession number. Bootstrap values are shown and the scale bar indicates the number of nucleotide substitutions per site

outermost cluster in the tree (subtree B), being an outgroup of the alleles in subtree A. Compared with the three sequences in subtree B, DQ889708 is closer to the alleles cloned in this study, located on a peripheral clade of A and D genome tribes.

Discussion

Phylogenetic relationships of the PPO alleles and the origin of common wheat

In the present study, three trees were generated based on different algorithms; nevertheless, topologies of the trees were highly similar to each other, implying the validity of the phylogenetic relationships. Only the relative positions of *Ppo-D1b* and DQ889708 were differed from the other three trees as mentioned above, which were probably due to the short sequence of DQ889708, making variable phylogenetic inferences when different algorithms were used. As expected, three partial PPO sequences, DQ889693, DQ889700 and DQ889706, were clustered together as an outgroup (subtree B) of the major group including A and D genome tribes and DQ889708 (subtree A). According to Massa et al. (2007), the three partial PPO sequences belonged to three out of four clusters, respectively, and the members of the fourth cluster (designated as cluster IV in Massa's paper) were all found in the subtree A constructed in this study, demonstrating that the subtree A is corresponding to Massa's group IV. Thus, we could conclude that the sequences in subree B are paralogs of the alleles in subtree A, while those within the subtree A are orthologous genes. In addition, the conclusion is also supported by the characteristics of the second introns. A *Stowaway* element reported previously by Massa et al. (2007) was inserted into DQ889708, and its putative footprints were found in other two PPO gene sequences, DQ889709 and DQ889710, corresponding to *Ppo-A1b* and *Ppo-A1e* identified in this study, respectively. Interestingly, the putative footprints characterized by the consensus sequence 5′TCCCT3′ were observed in the second introns from all the members within the subtree A, whereas not present in subtree B (Fig. 2, DQ889693 and DQ889700 did not harbor any introns). This result implied that the *Stowaway* element might have been inserted into the ancestor gene of the subtree A, and subsequently excised from *Ppo-A1* and *Ppo-D1*, leaving footprints in the gene sequences, whereas the element remained in DQ889708. Based on these facts, it is tempting to say that DQ889708 must be the ortholog of *Ppo-A1* and *Ppo-D1* on the chromosome 2BL, *Ppo-B1*, responsible for the QTL detected on this chromosome (Demeke et al. 2001; Watanabe et al. 2004). Based on these analyses, it could be come into conclusion that all of the alleles in the subtree A must be located on homoeologous group 2 chromosomes, associated with grain PPO activity.

The A genome of common and durum wheats was originated from *T. urartu* (Dvorak et al. 1993; Ciaffi et al. 2000; Huang et al. 2002; Petersen et al. 2006), which is in agreement with the results inferred from the structure of the cluster I (Fig. 3), where the allele from *T. urartu* was clustered with those from tetraploid and hexaploid wheats. However, an allele found in seven *T. durum* accessions (represented by *T. durum* DR8, Table 1), *Ppo-A1e*, was clustered with *T. monococcum*, sharing exactly the same allele with the latter. From this fact, it seems that the A genome of *T. durum* DR8 was originated from *T. monococcum*, however, this hypothesis is in conflict with the results from many other studies, which suggested that *T. urartu* is the A genome donor of tetraploid wheat species (Dvorak et al. 1993; Ciaffi et al. 2000; Huang et al. 2002; Petersen et al. 2006). Vardi (1973) demonstrated that gene exchange could occur between diploid and tetraploid wheat species by a triploid bridge, and therefore, the situation encountered here may be attributed to the result of a recent introgression event. A similar phenomenon was also observed by Zhang et al. (2002), in which the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA in *T. dicoccoides*, *T. monococcum* and *T. urartu* were analyzed, and the ITS repeat from *T. dicoccoides* is markedly similar to that from *T. monococcum*, other than that from *T. urartu*, showing the existence of gene exchange between diploid and tetraploid wheat species. Nevertheless, we cannot exclude another possibility, which is the existence of a *T. urartu* accession with *Ppo-A1e* allele, since we only sampled one *T. urartu* accession in this study.

A 191-bp fragment was found to be inserted into the first intron of *Ppo-A1b* from common wheat (Sun et al. 2005; He et al. 2007). Coincidently, the fragment was also found in *T. dicoccoides* DS4,

which shared the same allele with common wheat cv. Chinese Spring, showing clearly that the fragment in Chinese Spring was inherited from an ancestor closely related to *T. dicoccoides* DS4. On the other hand, however, the allele *Ppo-A1a* of common wheat showed high sequence identity with that of *T. dicoccoides* DS3 (*Ppo-A1f*), with only one SNP detected, suggesting that *Ppo-A1a* might have evolved from *Ppo-A1f*. Therefore, we can make an inference that more than one tetraploid emmer wheat participated in the formation of hexaploid common wheat. A similar suggestion was also proposed by Gu et al. (2004) when analyzing the A genome HMW-glutenin locus in common and durum wheats. In addition, Ozkan et al. (2005) identified two different genetic lineages within *T. dicoccoides* based on AFLP data, and also found the two lineages in *T. durum* and *T. dicoccon* lines. Furthermore, Haudry et al. (2007) investigated the nucleotide diversity at 21 loci and found two groups of common wheat, with each having its related tetraploid wheat lines. All these findings support the recurrent origin of common wheat, involving more than one tetraploid progenitor.

The D genome of common wheat derived from *Ae. tauschii* some 8,000 years ago (Huang et al. 2002). The results from Hammer (1980) implied that more than one *Ae. tauschii* lines participated in the formation of hexaploid wheat, and in the past decade, increasing molecular evidences supported this conclusion (Dvorak et al. 1998; Lelley et al. 2000; Caldwell et al. 2004; Giles and Brown 2006). In our previous study (He et al. 2007), two distinct alleles at *Ppo-D1* locus, *Ppo-D1a* and *Ppo-D1b*, were identified in common wheat cultivars, which were hypothesized to be inherited from different *Ae. tauschii* donors. In this study, *Ppo-D1a* showed great sequence identity with the allele from *Ae. tauschii* Ae38 (*Ppo-D1c*), with only one SNP detected, indicating that *Ppo-D1a* might have been derived from an *Ae. tauschii* donor closely related to *Ae. tauschii* Ae38. However, the ancestor of *Ppo-D1b* was not identified in this study, and as shown in the phylogenetic tree (Fig. 3), *Ppo-D1b* was located on the most outer layer of the A and D genome tribes, diverged not only from the alleles on the A genome, but also from the remaining alleles on the D genome. This may be attributed to the limited samples of *Ae. tauschii* accessions, omitting the *Ae. tauschii* lines with an allele related to the ancestor of *Ppo-D1b*. Another possible reason is that, like the example of *T. durum* DR8, *Ppo-D1b* was introduced into the D genome of common wheat from another species, which was not included in this study, by an introgression event.

Allelic variants of PPO genes and grain PPO activity

Grain PPO activity is an important trait both in common and durum wheats. The two PPO genes highly associated with PPO activity, *Ppo-A1* and *Ppo-D1*, and their allelic variations in common wheat were well characterized (Sun et al. 2005; He et al. 2007). In the present study, we identified seven new alleles of *Ppo-A1* and *Ppo-D1* in species related to common wheat, which might enable us to further understand the molecular mechanism of grain PPO activity.

Two alleles of *Ppo-A1* in common wheat, *Ppo-A1a* and *Ppo-A1b*, were proven to be associated with higher and lower PPO activities, respectively, and the molecular mechanism for the different phenotypic effects is probably due to the alternative splicing of the first intron, caused by a 191-bp insertion (Sun et al. 2005; He et al. 2007). Interestingly, the 191-bp insertion sequence was also found in *T. dicoccoides*, and the associations of its presence or absence in the species with PPO activity still need to be investigated. Fuerst et al. (2008) reported three *T. dicoccoides* lines with low PPO activity. Other studies, however, indicated *T. dicoccoides* lines with high PPO activity (Simeone et al. 2002; Watanabe et al. 2004), implying a wide range of phenotypic variation in this species, which may be associated with the 191-bp InDel.

Previous studies indicated that although many durum wheat lines had lower PPO activity, there are still a number of varieties that exhibited moderate to high PPO activity (Bernier and Howes 1994; Watanabe et al. 2004; Fuerst et al. 2008). Fuerst et al. (2008) suggested that the high PPO activity in *T. durum* varieties might be inherited from *T. monococcum* with high PPO activity in all accessions. As shown in the present study, two distinct alleles, *Ppo-A1e* and *Ppo-A1g*, were found in *T. durum* lines, and it is noteworthy that *Ppo-A1e* was probably introgressed from *T. monococcum*, which showed high PPO activity, thus the *T. durum* lines exhibiting high PPO activity could contain this allele. On the other hand, most *T. durum* varieties, represented by the cultivar Langdon, showed very low PPO activities (Bernier and Howes 1994;

Fuerst et al. 2008). The allele from *T. durum* 'Langdon' (*Ppo-A1g*) was analyzed in this study, however, only the upstream sequence of the allele was obtained, while its downstream sequence remains unknown, though many strategies such as Inverse PCR and Tail-PCR were tried. Considering the low PPO activity characterized in *T. durum* wheat line 'Langdon', the carrier of *Ppo-A1g*, the downstream sequence of this allele may have numerous mutations which result in a partial or even complete inactivation of this allele. Low sequence homology would also explain why the downstream sequence could not be amplified in this study, using primers designed based on the alleles from common wheat.

Two recent reports (Chang et al. 2007; Fuerst et al. 2008) indicate that *Ae. tauschii* may generally have moderate-to-low PPO activity. In the present study, we identified two alleles in this species, *Ppo-D1c* and *Ppo-D1d*, and the former shared high sequence identity with *Ppo-D1a* associated with low PPO activity in common wheat, implying that the allele *Ppo-D1c* could also be related to low PPO activity. *Ppo-D1d*, however, possessed a 73-bp deletion within the third exon (Fig. 1), leading to a premature translation termination. The two alleles may be associated with the low PPO activity in *Ae. tauschii*.

Acknowledgements The authors are very grateful to Dr. Lihui Li, Institute of Crop Science, CAAS, for his kindly providing the wheat related species. This study was supported by National Science Foundation of China (30771335), National 863 Programs (2006AA10Z1A7 and 2006AA100102), and International Collaboration Project from the Ministry of Agriculture (2006-G2).

References

Anderson JV, Morris CF (2001) An improved whole-seed assay for screening wheat germplasm for polyphenol oxidase activity. Crop Sci 41:1697–1705

Anderson JV, Morris CF (2003) Purification and analysis of wheat grain polyphenol oxidase protein. Cereal Chem 80:135–143

Baik BK, Czuchajowska Z, Pomeranz Y (1995) Discoloration of dough for oriental noodles. Cereal Chem 72:198–205

Bernier AM, Howes NK (1994) Quantification of variation in tyrosinase activity among durum and common wheat cultivars. J Cereal Sci 19:157–159

Boss PK, Gardner RC, Janssen BJ, Ross GS (1995) An apple polyphenol oxidase cDNA is up-regulated in wounded tissues. Plant Mol Biol 27:429–433

Caldwell KS, Dvorak J, Lagudah ES, Akhunov E, Luo MC, Wolters P et al (2004) Sequence polymorphism in polyploid wheat and their D-genome ancestor. Genetics 167:941–947

Chang C, Zhang HP, Xu J, You MS, Li BY, Liu GT (2007) Variation in two PPO genes associated with polyphenol oxidase activity in seeds of common wheat. Euphytica 154:181–193

Ciaffi M, Dominici L, Umana E, Tanzarella OA, Porceddu E (2000) Restriction fragment length polymorphism (RFLP) for protein disulfide isomerase (PDI) gene sequences in *Triticum* and *Aegilops* species. Theor Appl Genet 101:220–226

Constabel CP, Yip L, Patton JJ, Christopher ME (2000) Polyphenol oxidase from hybrid poplar: Cloning and expression in response to wounding and herbivory. Plant Physiol 124:285–295

Demeke T, Morris CF, Campbell KG, King GE, Anderson JA, Chang H (2001) Wheat polyphenol oxidase: distribution and genetic mapping in three inbred line populations. Crop Sci 41:1750–1757

Dexter JE, Preston KR, Matsuo RR, Tipples KH (1984) Development of a high extraction flour for the GRL Pilot Mill to evaluate Canadian wheat potential for the Chinese market. Can Inst Food Sci Technol 14:253–259

Dvorak J, di Terlizzi P, Zhang HB, Resta P (1993) The evolution of polyploid wheats: identification of the A genome donor species. Genome 36:21–31

Dvorak J, Luo MC, Yang ZL, Zhang HB (1998) The structure of the *Aegilops tauschii* gene pool and the evolution of hexaploid wheat. Theor Appl Genet 97:657–670

Feillet P, Autran JC, Icard-Vernière C (2000) Pasta brownness: an assessment. J Cereal Sci 32:215–233

Feuillet C, Penger A, Gellner K, Mast A, Keller B (2001) Molecular evolution of receptor-like kinase genes in hexaploid wheat. Independent evolution of orthologs after polyploidization and mechanisms of local rearrangements at paralogous loci. Plant Physiol 125:1304–1313

Fuerst EP, Anderson JV, Morris CF (2006) Delineating the role of polyphenol oxidase in the darkening of alkaline wheat noodles. J Agric Food Chem 54:2378–2384

Fuerst EP, Xu SS, Beecher B (2008) Genetic characterization of kernel polyphenol oxidase in wheat and related species. J Cereal Sci. doi:10.1016/j.jcs.2007.10.003

Giles RJ, Brown TA (2006) *GluDy* allele variations in *Aegilops tauschii* and *Triticum aestivum*: implications for the origins of hexaploid wheats. Theor Appl Genet 112:1563–1572

Gill BS, Friebe B (2002) Cytogenetics, phylogeny and evolution of cultivated wheats. In: Curtis BC, Rajaram S, Gómez Macpherson H (eds) Bread wheat: improvement and production. Food and Agriculture Organization of the United Nations, Rome, pp 71–88

Gooding PS, Bird C, Robinson SP (2001) Molecular cloning and characterization of banana fruit polyphenol oxidase. Planta 213:748–757

Gu YQ, Coleman-Derr D, Kong XY, Anderson OD (2004) Rapid genome evolution revealed by comparative sequence analysis of orthologous regions from four *Triticeae* genomes. Plant Physiol 135:459–470

Hammer K (1980) Vorarbeiten zur monographischen Darstellung von Wildpflanzensortimenten: *Aegilops* L. Kulturpflanze 28:33–180

Haruta M, Pederson JA, Constabel CP (2001) Polyphenol oxidase and herbivore defense in trembling aspen (*Populus tremuloides*): cDNA cloning, expression, and potential substrates. Physiol Plant 112:552–558

Haudry A, Cenci A, Ravel C, Bataillon T, Brunel D, Poncet C et al (2007) Grinding up wheat: a massive loss of nucleotide diversity since domestication. Mol Biol Evol 24:1506–1517

He XY, He ZH, Zhang LP, Sun DJ, Morris CF, Fuerst EP et al (2007) Allelic variation of polyphenol oxidase (PPO) genes located on chromosomes 2A and 2D and development of functional markers for the PPO genes in common wheat. Theor Appl Genet 115:47–58

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R et al (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci USA 99:8133–8138

Jauhar PP (2007) Meiotic restitution in wheat polyhaploids (amphihaploids): a potent evolutionary force. J Hered 98(2):188–193

Jimenez M, Dubcovsky J (1999) Chromosome location of genes affecting polyphenol oxidase activity in seeds of common and durum wheat. Plant Breed 118:395–398

Jukanti AK, Bruckner PL, Fischer AM (2004) Evaluation of wheat polyphenol oxidase genes. Cereal Chem 81:481–485

Jukanti AK, Bruckner PL, Fischer AM (2006) Molecular and biochemical characterisation of polyphenol oxidases in developing kernels and senescing leaves of wheat (*Triticum aestivum*). Funct Plant Biol 33:685–696

Kruger JE (1976) Changes in the polyphenol oxidases of wheat during kernel growth and maturation. Cereal Chem 53:201–213

Kruger JE, Matsuo RR, Preston K (1992) A comparison of methods for the prediction of Cantonese noodle colour. Can J Plant Sci 72:1021–1029

Kruger JE, Hatcher DW, DePauw R (1994) A whole seed assay for polyphenol oxidase in Canadian prairie spring wheats and its usefulness as a measure of noodle darkening. Cereal Chem 71:324–326

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163

Lagudah ES, Appels R, McNeil D (1991) The *Nor-D3* locus of *Triticum tauschii*: natural variation and genetic linkage to markers in chromosome 5. Genome 34:387–395

Lelley T, Stachel M, Grausgruber H, Vollmann J (2000) Analysis of relationships between *Ae. tauschii* and the D genome of wheat utilizing microsatellites. Genome 43:661–668

Luo MC, Yang ZL, You FM, Kawahara T, Waines JG, Dvorak J (2007) The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. Theor Appl Genet 114:947–959

Mares DJ, Campbell AW (2001) Mapping components of flour and noodle colour in Australian wheat. Aust J Agric Res 52:1297–1309

Massa AN, Beecher B, Morris CF (2007) Polyphenol oxidase (PPO) in wheat and wild relatives: molecular evidence for a multigene family. Theor Appl Genet 114:1239–1247

McCallum JA, Walker JRL (1990) *O*-diphenol oxidase activity, phenolic content and colour of new Zealand wheats, flours and milling streams. J Cereal Sci 12:83–96

Newmann SM, Eannetta NT, Yu H, Prince JP, de Vicente CM, Tanksley SD et al (1993) Organization of the tomato polyphenol oxidase gene family. Plant Mol Biol 21:1035–1051

Ozkan H, Brandolini A, Pozzi C, Effgen S, Wunder J, Salamini F (2005) A reconsideration of the domestication geography of tetraploid wheats. Theor Appl Genet 110:1052–1060

Petersen G, Seberg O, Yde M, Berthelsen K (2006) Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). Mol Phylogenet Evol 39:70–82

Raman R, Raman H, Johnstone K, Lisle C, Smith A, Matin P et al (2005) Genetic and in silico comparative mapping of the polyphenol oxidase gene in bread wheat (*Triticum aestivum* L.). Funct Integr Genomics 5:185–200

Salamini F, Özkan H, Brandolini A, Schäfer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the Near East. Nat Rev Genet 3:429–441

Simeone R, Pasqualone A, Clodoveo ML, Blanco A (2002) Genetic mapping of polyphenol oxidase in tetraploid wheat. Cell Mol Biol Lett 7:763–769

Sullivan ML, Hatfield RD, Thoma SL, Samac DA (2004) Cloning and characterization of red clover polyphenol oxidase cDNAs and expression of active protein in *Escherichia coli* and transgenic Alfalfa. Plant Physiol 136:3234–3244

Sun DJ, He ZH, Xia XC, Zhang LP, Morris CF, Appels R et al (2005) A novel STS marker for polyphenol oxidase activity in bread wheat. Mol Breed 16:209–218

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Thygesen PW, Dry IB, Robinson SP (1995) Polyphenol oxidase in potato. Plant Physiol 109:525–531

Vardi A (1973) Introgression between different ploidy levels in the wheat group. In: Sears ER, Sears LMS (eds) Proceedings of the 4th international wheat genetics symposium, Columbia, Missouri, USA, pp 131–141

Watanabe N, Takeuchi A, Nakayama A (2004) Inheritance and chromosomal location of the homoeologous genes affecting phenol colour reaction of kernels in durum wheat. Euphytica 139:87–93

Zhang W, Qu LJ, Gu H, Gao W, Liu M, Chen J et al (2002) Studies on the origin and evolution of tetraploid wheats based on the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA. Theor Appl Genet 104:1099–1106